Title:	**Help to make Optimal Decisions from Complex Data**
Author(s):	Greg Wilson
Submitted to:	
	http://lib-www.lanl.gov/la-pubs/00818562.pdf

Help to make Optimal Decisions from Complex Data

Greg Wilson
D-1, Statistical Sciences Group
Decision Applications Division
Los Alamos National Laboratory

The Statistical Science Group at DOE's Los Alamos National Laboratory (LANL) has been working the last two years to understand how decision makers in the oil and gas industry can make better decisions by understanding the uncertainties associated with complex systems and data. Under a cooperative research agreement with the Petrochemical Open Software Corporation (POSC), Pacific Operators Offshore, Inc. (POOI), and Belden & Blake Corporation, and utilizing funding from DOE's Advanced Computational Technology Initiative, the Statistical Sciences Group has been developing methods to characterize complex systems, identify and integrate diverse types of data, and make calculations to predict total system performance. Systems may be thought of as man-made (like refineries) or more natural (like reservoirs). With an understanding of how they system is likely to "perform," decision makers can factor that information into higher-level decisions or think about how to allocate resources to improve the performance of the system. Decision makers can also use the LANL methods to run "what if" scenarios to see how changes in the system may affect performance.

Development of these methods has been a collaboration between social scientists who know how to elicit information from technical experts, statisticians who know how to characterize uncertainties, and software engineers who are beginning to automate these methods into more standardized tools. At the beginning of a project to model a particular system, social scientists meet with the community of experts who own the system. Using ethnographic methods, the social scientists elicit from the experts a model of the system that represents the key components, features, factors, and relationships. The goal is to capture the pertinent knowledge that the experts have about the system. This "knowledge model" is initially represented graphically as a diagram of boxes that show components, features, and factors connected by lines and arrows that show relationships.

As the knowledge model becomes robust enough that everyone feels it is a good characterization of the complex system, the social scientists and statisticians begin exploring what kind of data is available to fill each of the boxes and how the relationships between the boxes can be understood in mathematical terms. For each box, there may be traditional test data, there may be data on a related system, there may be computer model output, or there may be no data at all. In the absence of data, the social scientists employ methods to elicit and quantify the expert judgment of the people with the most knowledge about that box. Formally utilizing such expertise may be untraditional, but often the experts provide fantastic data based on their understanding of how that piece was designed to work. And certainly, taking advantage of what you do know is better than leaving a gaping hole in your model.

In many cases, there exist multiple types of data for each of the boxes. Part of the challenge in the development of these methods has been understanding how to integrate these diverse types of data to get a accurate picture of the reliability of each piece of the

system, and then to integrate all of those pieces into an estimate of the reliability and performance of the system as a whole. As these calculations and evaluations take place, the knowledge model is morphed into a graphical statistical model that is mathematically and computationally tractable. Each system, though, presents its own peculiarities and challenges, and developing appropriate statistical models that accurately integrate the data and then accurately integrate the constituent uncertainties is no small problem. All areas of this research are expanding the limits of statistical and social science methodologies.

The Statistical Sciences Group is eager to continue address the ongoing problems that the oil and gas industry faces, but the methods under development are applicable to many other types of problems. These methods are being successfully applied in the nuclear weapons program at LANL, in several Department of Defense projects, and with other industry collaborators. In fact, similar work with Delphi Automotive garnered the group an R&D 100 award in 1999. The Statistical Science Group is always looking for new industrial partners because new problems always offer the possibility of new answers and new ways to get at those answers. If you are interested in finding out more about the work being done in this area at LANL contact Greg Wilson, gdwilson@lanl.gov.